Euler Project - Problem 6 - Sum Square Difference

Task and corollaries

The sum of the squares of the first ten natural numbers is,

$$1^2 + 2^2 + \ldots + 10^2 = 385.$$

The square of the sum of the first ten natural numbers is,

$$(1 + 2 + \ldots + 10)^2 = 55^2 = 3025.$$

Hence the difference between the sum of the squares of the first ten natural numbers and the square of the sum is $3025 - 385 = 2640$.

Find the difference between the sum of the squares of the first one hundred natural numbers and the square of the sum.

Source: $https://projecteuler.net/problem=6$

The needed corollaries for this proof can be studied in Appendix A.
Assumed is knowledge of summation laws and arithmetic series.

Finding the pattern

To better describe the patterns, let's assign letters to the specific natural numbers, starting with $a = 1$, $b = 2$, and so on.

Demonstrated on $N = 2$:

$$\mathbf{N} = 2$$
$$1^2 + 2^2 \rightarrow a^2 + b^2$$
$$= 5$$
$$(1 + 2)^2 \rightarrow (a + b)^2 = a^2 + 2ab + b^2$$
$$= 9$$

$$\Delta = 4$$

Notice the highlighted terms appear in both expressions. As a consequence of the Binomial expansion the first and last terms will always appear squared. Clearly, the difference consists purely of the other terms in the expansion.

Let's shift our focus on the expansion of the second expression (sum squared) only:

$$\mathbf{N} = 3$$
$$(a + b + c)^2 = (a^2 + ab + ac) + (ba + b^2 + bc) + (ca + cb + c^2)$$
$$= a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$$

$$\mathbf{N} = 4$$
$$(a + b + c + d)^2 = \ldots + 2ab + 2ac + 2ad + 2bc + 2bd + 2cd$$

The FOIL expansion gives a hint as to what's going on - after summing the terms, we may notice that for any N, the difference is the product of the terms times two, for every pair of terms, each multiplied with each other exactly once.

Generally, for N, where a=1, b=2 etc. until N:

1. In the inner brackets, we [summarize N-1 terms, starting with a, up until N-1]. We will call each term "i"
2. We multiply those by the [sum of N-1 terms, starting with i+1, up until N]. We will call each term "j"

In other words, starting with a, we do:
$a \cdot b + a \cdot c + a \cdot d + ... + a \cdot N$

And repeat for each number, increasing by 1 up until N-1, e.g. the next sequence would be:
$b \cdot c + b \cdot d + ... + b \cdot N$

We then sum all these sequences, and multiply the result by 2. Notice we described the observed pattern:
$2[a(b + c + d) + b(c + d) + c(d) + ... + (N - 1)(N)]$

Deriving the formula

$$2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} i \cdot j \qquad\qquad \text{rewritten as a double sum}$$

$$2 \sum_{i=1}^{N-1} i \sum_{j=i+1}^{N} j \qquad\qquad \text{extract independent term}$$

$$2 \sum_{i=1}^{N-1} i \left(\sum_{j=1}^{N} j - \sum_{j=1}^{i} j \right) \qquad\qquad \text{by } \textbf{Lemma 1}$$

$$2 \sum_{i=1}^{N-1} i \left(N \frac{1 + N}{2} - i \frac{1 + i}{2} \right) \qquad\qquad \text{by } \textbf{Lemma 2.1}$$

$$2 \sum_{i=1}^{N-1} \frac{i}{2} [N(1 + N) - i(1 + i)] \qquad\qquad \text{factor out } \frac{1}{2}$$

$$\sum_{i=1}^{N-1} i [N(1 + N) - i(1 + i)] \qquad\qquad c \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} c a_i$$

$$\sum_{i=1}^{N-1} i [N^2 + N - (i^2 + i)] \qquad\qquad \text{distribute all brackets}$$

$$\sum_{i=1}^{N-1} (iN^2 + iN - i^3 - i^2)$$

$$N^2 \sum_{i=1}^{N-1} i + N \sum_{i=1}^{N-1} i - \sum_{i=1}^{N-1} i^3 - \sum_{i=1}^{N-1} i^2 \qquad\qquad \sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$$

$$N^2 \frac{(N-1)N}{2} + N \frac{N(N-1)}{2} - \frac{(N-1)^2(N^2)}{4} - \frac{N(N-1)(2N-1)}{6} \qquad\qquad \text{by } \textbf{Lemmas 1.1, 1.2, 1.3}$$

$$N(N-1) \left(\frac{N^2}{2} + \frac{N}{2} - \frac{(N-1)N}{4} - \frac{2N-1}{6} \right) \qquad\qquad \text{factor out } N(N-1)$$

$$N(N-1) \left(\frac{6N^2}{12} + \frac{6N}{12} - \frac{3N^2 - 3N}{12} - \frac{4N - 2}{12} \right) \qquad\qquad \text{combine fractions}$$

$$N(N-1)\frac{3N^2+5N+2}{12}$$

$$N(N-1)\frac{(3N+2)(N+1)}{12} \qquad \text{factor the numerator}$$

$$N\frac{(N^2-1)(3N+2))}{12} \qquad (a+b)(a-b)=a^2-b^2$$

$$N\frac{3N^3+2N^2-3N-2}{12}$$

$$\frac{3N^4+2N^3-3N^2-2N}{12} \qquad \text{distribute and simplify}$$

And we arrive at a closed form formula for the sum square difference for any N.

Answering the question $(N=100)$

> Find the difference between the sum of the squares of the first **one hundred** natural numbers and the square of the sum.

Using the formula derived earlier:

$$\frac{3(100)^4+2(100)^3-3(100)^2-2(100)}{12}$$

$$\frac{3(100000000)+2(1000000)-3(10000)-2(100)}{12}$$

$$\frac{300000000+2000000-30000-200}{12}$$

$$\frac{302000000-30200}{12}$$

$$\frac{301969800}{12}$$

$$= \mathbf{25164150}$$

Euler Project - Problem 6 - Appendix A

Corollaries & Identities

Hello there!